Exercise 1:

This week our research group is still busy with the task of deciding whether a patient should be admitted to the hospital. However, based on the findings of the last weeks, it is questioned whether a linear classifier can actually adequately separate the patients into two groups.

1) **Explain why neural networks (NNs) can be useful in such a situation and how they relate to the concept of representation learning.**

Now that we know why NNs can be useful, researcher Holger wonders how such an NN actually works. Researcher Stefanie states that an *input layer*, *neurons, hidden layers* and an *output layer* are the central components of a NN.
(Tip: A small sketch of an NN might be helpful for the next tasks)

2) **Let's start with the input layer. Explain its role in the network and how it interacts with subsequent components.**

After the input layer, an arbitrary number of so-called hidden layers follows. Each of these layers consists of an arbitrary number of neurons.

3) **Explain why these layers are considered to be "hidden".**

4) **Each neuron in a hidden layer performs a 2-step computation. Explain what these two steps are and where the so-called *weights* come into play here.**

The last component of an NN is the so-called output layer, which again consists of a number of neurons.

5) **Explain how the number of neurons and the activation function, which is used in these/this neuron(s) of the output layer, relate to the respective ML problem. For example, consider how the assumed distribution of the target variable in a classification/regression problem affects the choice of function and the number of neurons.**

Now that we know the essential parts of NNs, we can apply this knowledge to our already familiar task, whether a patient should be admitted to the hospital. As you already know, this is a binary classification task with target space $\mathcal{Y} = \{0, 1\}$, with $y = 1$ if the patient requires intensive care and $y = 0$ if not. The feature space is: $\mathcal{X} = (\mathbb{R}_0^+)^3$, with $\mathbf{x}^{(i)} = (x_{age},\ x_{blood\,pressure},\ x_{weight})^{(i)} \in \mathcal{X}$ for $i = 1, 2, \ldots, n$ observations.

6) **Draw a sketch of a suitable NN for this task. It should consist of two hidden layers, each containing three neurons. In addition, each of the features from the feature space should be used as input. Furthermore, specify the activation function used in the output layer.**

7) **Based on your sketch, figure out how many parameters need to be estimated during training.**